STATISTICAL RESEARCH REPORT
Institute of Mathematics
University of Oslo

No 4
1975

A GENERAL PRINCIPLE FOR A MULTIPLE COMPARISON PROCEDURE FOR LINEAR CONTRASTS, AND SOME APPLICATIONS.

by

Jan F. Bjørnstad.

ABSTRACT

In this paper multiple comparison methods for linear contrasts of a set of parameters $\theta_1, \ldots, \theta_k$ is considered in a general context. A linear contrast is a linear function $\sum_i c_i \theta_i$ with $\sum_i c_i = 0$. It is shown that a test for the hypothesis $\theta_1 = \ldots = \theta_k$ rejects if and only if a proposed procedure for linear contrasts states at least one contrast greater than zero. Within a set of very wide asymptotic assumptions it is shown that the proposed procedure for contrasts satisfy that the probability of at least one false statement is asymptotically less than or equal to a fixed ϵ. Also simultaneous confidence intervals for all contrasts are given in the asymptotic case. The general result is applied to comparison of contingency tables, the one-way layout in the analysis of variance with unknown group variances, and to comparison of binomial probabilities. The usual "analysis of variance" situation is also considered.

Key words: linear contrast, multiple test procedure, simultaneous significance level.

CONTENTS

1. Introduction.

We are given a set of unknown parameters $\theta_1, \ldots, \theta_k$ which we want to compare. Our task is to find out which of the linear contrasts $\sum_i c_i \theta_i$, $\sum_i c_i = 0$, that are greater than zero. The difference $\theta_i - \theta_j$ is of course a contrast, but we will not restrict ourselves to such comparisons only. It could be that other contrasts are equally or more interesting to us. Therefore we consider a comparison rule which allow us to "look at" all contrasts.

A general principle for constructing a comparison rule for all linear contrasts will be given. Under very wide asymptotic assumptions it is shown that a comparison rule based on the principle will satisfy that the probability of at least one false statement among all contrasts is asymptotically less than or equal to a fixed quantity ϵ . In special cases, as in the analysis of variance, a comparison rule based on the principle will have a simultaneous level ϵ also for a finite number of observations. As a matter of fact it is seen that the Scheffé method for comparing contrasts in the "analysis of variance model" follows from our principle. We also consider the one-way layout of analysis of variance where the group variances are unknown. Spjøtvoll, [5] , has given a procedure for general linear functions of $\theta_1, \ldots, \theta_k$ in this case. This procedure has a natural modification to linear contrast, which we show is valid asymptotically, by applying our general asymptotic result mentioned above. Another application of this general result is to comparison of contingency tables.

As a follow-up of the author's paper [2] , we give a method for judging all contrasts when comparing independent two-way contingency tables by use of measures of association. In [2] the author proposed a method for general linear functions, and the rule for linear contrasts turns out to be a natural modification of that method.

The most important aspect of our general comparison method is that it has the property of stating at least one significant contrast if and only if a relating test-statistic for the homogenity hypothesis

$$H: \theta_1 = \ldots = \theta_k \, . \tag{1}$$

is significant, i.e. the hypothesis is rejected. This is a very desirable property of a comparison rule for contrast, since

$$\sum_i c_i \theta_i = 0 \quad \text{for all contrast}$$

if and only if

$$\theta_1 = \ldots = \theta_k \, .$$

If then we reject the homogenity hypothesis then we also have to state at least one contrast different from zero, and vice versa. This is exactly what our proposed method does.

In the next section we state our principle of equivalence which is purely an algebraic result.

2. A principle for constructing methods for linear contrasts.

The usual notation of vector space algebra is used, so that for a matrix A, A' denotes the transpose of A. If V is a vector space then by $d(V)$ we mean the dimension of V. $\mathbb{R}^k$ denotes the k-dimensional euclidian space with real components.

First we state an algebraic result which will be the fundament of our theory. Thereafter we state the principle.

LEMMA 1. *Let $y = (y_1,\ldots,y_k)'$ be a k-dimensional vector and let B be a $k \times t$ $(k>t)$-dimensional matrix with rank equal to r, satisfying $B'y = 0$. Let further $\mathcal{L} = \{h \in \mathbb{R}^k \mid h'B = 0\}$. Then*

$$\begin{array}{c} h'v \le \sqrt{z}\sqrt{h'h} \quad \forall h \in \mathcal{L} \\ \Updownarrow \\ y'y \le z \end{array} \qquad (2)$$

Proof.

Let V_r be the subspace in $\mathbb{R}^k$ generated by the columns in B. Then $d(V_r) = r$, since rank $(B) = r$. Let $H_1,\ldots,H_r$ be an orthonormal basis for V_r. Put $H = (H_1,\ldots,H_r)$. Then $H'H = I$, the identity matrix. Let now G be a $k \times (k-r)$ dimensional matrix consisting of orthonormal columns so that $K = (G,H)$ constitutes an orthonormal basis of $\mathbb{R}^k$. Then $K'K = I$ and $KK' = I$.

Let for all $h \in \mathcal{H}$: $d = K'h$, and let $\mathcal{D} = \{d \in \mathbb{R}^k : d = K'h, h \in \mathcal{H}\}$.
Let $v = K'y$. $h = K \circ d$ for all $h \in \mathcal{H}$ and $y = K \circ v$.
Hence

$$h'y = d'v \quad \text{and} \quad y'y = v'v \ .$$

Since $h \perp B$ (i.e. orthogonal to the columns in B), $h \perp V_r$ implying that $h \perp H$, i.e. $h'H = 0$ for $h \in \mathcal{H}$. This gives for $h \in \mathcal{H}$:

$$0 = h'H = d'K'H = (d_{k-r+1}, \ldots, d_k)$$

$$\Rightarrow \quad d_{k-r+1} = \ldots = d_k = 0 \ .$$

Let $f_h = h'y$. Then $f_h = d'v = \sum_{i=1}^{k-r} d_i v_i$,

$$\text{and} \quad h'h = d'd = \sum_{i=1}^{k-r} d_i^2$$

Further we have that

$$v = \begin{pmatrix} G'y \\ H'y \end{pmatrix} = \begin{pmatrix} G'y \\ 0 \end{pmatrix} \ , \quad \text{since} \quad B'y = 0 \quad \text{and therefore} \quad H'y = 0 \ .$$

This implies

$$v_{k-r+1} = \ldots = v_k = 0 \quad \text{and} \quad y'y = \sum_{i=1}^{k-r} v_i^2 \ .$$

Now $f_h \leq \sqrt{h'h} \quad \forall h \in \mathcal{H}$

$\wedge \| \vee$ (1)

$$\sum_{i=1}^{k-r} d_i v_i \leq \sqrt{z} \sqrt{\sum_{i=1}^{k-r} d_i^2} \quad \forall d \in \mathcal{D}$$

$\wedge \| \vee$ (2)

$$\sum_{i=1}^{k-r} v_i^2 \leq z \ .$$

(1) follows from the definition of $\mathfrak{D}$.

(2) follows from Schwartz inequality which states that

$$(\sum_{i=1}^{k-r} d_i v_i)^2 \leq \sum_{i=1}^{k-r} d_i^2 \cdot \sum_{i=1}^{k-r} v_i^2 \quad .$$

Hence

$$\sum_{i=1}^{k-r} v_i^2 \leq z \iff \sum_{i=1}^{k-r} d_i^2 \sum_{i=1}^{k-r} v_i^2 \leq z \sum_{i=1}^{k-r} d_i^2 \implies$$

$$(\sum_{i=1}^{k-r} d_i v_i)^2 \leq z \sum_{i=1}^{k-r} d_i^2 \quad , \quad \forall d \in \mathfrak{D} \, .$$

The other way: Assume

$$\sum_{i=1}^{k-r} d_i v_i \leq \sqrt{z}\sqrt{\sum_{i=1}^{k-r} d_i^2} \quad \forall d \in \mathfrak{D} \, .$$

Now $v \in \mathfrak{D}$, since $v = K'y$ and $y \in \mathcal{H}$. Hence

$$\sum_{i=1}^{k-r} v_i^2 \leq \sqrt{z}\sqrt{\sum_{i=1}^{k-r} v_i^2} \iff \sum_{i=1}^{k-r} v_i^2 \leq z \quad .$$

Now, since $y'y = \sum_{i=1}^{k-r} v_i^2$ the result follows.

Q.E.D.

To be able to state the principle of equivalence we introduce $\hat{\theta}_1,\ldots,\hat{\theta}_k$ as estimators of $\theta_1,\ldots,\theta_k$ respectively. Further $W_1,\ldots,W_k$ is a set of positive weights to $\hat{\theta}_1,\ldots,\hat{\theta}_k$ respectively, which can be either random variables or just known numbers. Now we want to estimate the joint value θ of $\theta_1,\ldots,\theta_k$ under the homogenity hypothesis (1) . Our estimation method is then to minimize

$$Q = \sum_{i=1}^{k} \frac{(\hat{\theta}_i - \theta)^2}{W_i} \qquad (3)$$

The estimator minimizing θ is denoted by $\bar{\theta}$. I.e. $\bar{\theta}$ satisfies

$$\sum_{i=1}^{k} \frac{(\hat{\theta}_i - \bar{\theta})}{W_i} = 0$$

This gives

$$\bar{\theta} = \sum_{i=1}^{k} \frac{\hat{\theta}_i}{W_i} / \left(\sum_{i=1}^{k} W_i^{-1} \right) \tag{4}$$

Let

$$U = \min_{H} Q = \sum_{i=1}^{k} \frac{(\hat{\theta}_i - \bar{\theta})^2}{W_i} \tag{5}$$

U is now a natural test-statistic for the hypothesis (1). The equivalence with a comparison procedure for contrasts follows in the next result.

THEOREM 1.

Let $\bar{\theta}$ and U be as defined in (4) and (5). Then

$$\begin{array}{c} U \leq z \\ \wedge \\ \| \\ \vee \end{array} \tag{6}$$

$$\sum_i c_i \hat{\theta}_i \leq \sqrt{z} \sqrt{\sum_i W_i c_i^2} \quad \text{for all} \quad c = (c_1, \ldots, c_k)' \text{ such that } \sum_i c_i = 0 .$$

Proof.

Let (t,w) be observed values of $(\hat{\theta}, W) = (\hat{\theta}_1, \ldots, \hat{\theta}_k, W_1, \ldots, W_k)$. Let $\bar{t}$ denote the corresponding value of $\bar{\theta}$.

Let further A be the event

$$A = \{(t,w) \mid \sum_i w_i^{-1}(t_i-\bar{t})^2 \leq z\} ,$$

and let

$$B = \{(t,w) \mid \sum_i c_i t_i \leq \sqrt{z}\sqrt{\sum_i w_i c_i^2} \text{ , all } c \text{ with } \Sigma c_i = 0\} .$$

We shall show that A = B or rather

$$\sum_i w_i^{-1}(t_i-\bar{t})^2 \leq z \iff \sum_i c_i t_i \leq \sqrt{z}\sqrt{\sum_i w_i c_i^2} \text{ , } \forall c \text{ , } \sum_i c_i = 0 .$$

Let now $y = (y_1,\ldots,y_k)'$ where $y_i = \sqrt{w_i^{-1}}(t_i-\bar{t})$ such that

$$U = y'y .$$

Put $B = (\sqrt{w_1^{-1}},\ldots,\sqrt{w_k^{-1}})'$. In this case t=r=1 in lemma 1.

Then

$$B'y = \sum_{i=1}^{k} \frac{t_i}{w_i} - \bar{t}\sum_{i=1}^{k} \frac{1}{w_i} = 0 .$$

Let $\mathcal{H} = \{h \in \mathbb{R}^k : h'B = 0\}$. $\mathcal{H}$ is a supspace of $\mathbb{R}^k$.

Let now $\mathcal{H}_1 = \{h \in \mathbb{R}^k : h_i = c_i\sqrt{w_i} \text{ , } \sum_i c_i = 0\}$. Here $h = (h_1,\ldots,h_k)'$.

It will now be shown that $\mathcal{H} = \mathcal{H}_1$.

Let us assume that $h \in \mathcal{H}_1 \Rightarrow h'B = \sum_i c_i = 0$. Hence $\mathcal{H}_1 \subset \mathcal{H}$.

$\mathcal{H}$ is the set of vectors orthogonal to the space spanned by B .

Then $d(\mathcal{H}) + \text{rank } (B) = k \Rightarrow d(\mathcal{H}) = k - 1$.

We return to $\mathcal{H}_1$. A basis for $\mathcal{H}_1$ is given by

$$e_1 = (\sqrt{w_1} \text{ , } 0,\ldots,0,-\sqrt{w_k})'$$

$$e_i = (0,\ldots,0,\sqrt{w_i},\ldots 0,-\sqrt{w_k})'$$

$$e_{k-1} = (0,\ldots,0,\sqrt{w_{k-1}},-\sqrt{w_k})'$$

Hence, $d(\mathcal{H}_1) = k-1$, giving $\mathcal{H}_1 = \mathcal{H}$.

We now have from lemma 1:

$$U \leq z$$
$$\wedge \quad \| \quad \vee$$
$$\sum_i h_i y_i \leq \sqrt{z}\sqrt{\sum_i h_i^2} \quad \forall h \in \mathcal{H}_1 .$$
$$\wedge \quad \| \quad \vee$$
$$\sum_i c_i (t_i - \bar{t}) \leq \sqrt{z}\sqrt{\sum_i c_i^2 w_i^2} \quad \forall c,\ \sum_i c_i = 0$$
$$\wedge \quad \| \quad \vee$$
$$\sum_i c_i t_i \leq \sqrt{z}\sqrt{\sum_i c_i^2 w_i^2} \quad \forall c,\ \sum_i c_i = 0 .$$

Q.E.D.

The weights $W_1,\ldots,W_k$ will usually be estimated variances of $\hat{\theta}_1,\ldots,\hat{\theta}_k$, such that $\sum_i W_i c_i^2$ is an estimator of the variance of $\sum_i c_i \hat{\theta}_i$. We therefore use the notation

$$\hat{\sigma}^2_{c,\hat{\theta}} = \sum_i W_i c_i^2 . \qquad (7)$$

In many cases with W_i as estimated variance of $\hat{\theta}_i$, U will have an exactly or asymptotically known distribution under the hypothesis (1) . Then the rejection region $\{U > z\}$, where z is the upper ε-fractile of this distribution, will be a test for H , and we have then obtained from theorem 1 a comparison rule with a simultaneous significance level exactly or asymptotically equal to ε under the homogenity hypothesis. Next we have to show that this is valid under any set of values of the parameters. This will be done later in a general way for the asymptotic case. First, however, we show that the equivalence of the S-method for linear contrasts and the F-test of homogenity in the "analysis of variance" case is a special case of theorem 1.

3. The "analysis of variance" situation. Normal case.

The model is as follows. $\hat{\theta}_1,\ldots,\hat{\theta}_k$ are independent and normally distributed with $E\hat{\theta}_i = \theta_i$ and $\operatorname{var}\hat{\theta}_i = b_i\sigma^2$ where b_i is known and σ^2 is unknown. Further S^2 is an estimator of σ^2 with the property that $\nu S^2/\sigma^2$ is chi-square distributed with ν degrees of freedom. $S^2, \hat{\theta}_1,\ldots,\hat{\theta}_k$ are independent.

This model covers many of the most interesting cases in the analysis of variance. We will mention some of them.

(i) One-way layout with equal group variances.

Model: X_{ij}, $i=1,\ldots,k$, $j=1,\ldots,n_i$ are independent and normally distributed with $EX_{ij} = \theta_i$ and $\operatorname{var} X_{ij} = \sigma^2$. Here is $\hat{\theta}_i = \bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$, so that $b_i = n_i^{-1}$. Let $n = \sum_{i=1}^{k} n_i$. The estimator of σ^2 is

$$S^2 = \frac{1}{n-k}\sum_{i,j}(X_{ij}-\bar{X}_i)^2 \quad \text{with} \quad \nu = n-k .$$

(ii) Two-way layout with equal numbers in the cells.

(a) Without interaction.

Model: X_{ij}, $i=1,\ldots,r$, $j=1,\ldots,s$ are independent and normally distributed.

$$EX_{ij} = \mu + \alpha_i + \beta_j \quad \text{where} \quad \sum_i \alpha_i = \sum_j \beta_j = 0, \quad \operatorname{var} X_{ij} = \sigma^2 .$$

Here

$$S^2 = \frac{1}{(r-1)(s-1)}\sum_{i,j}(X_{ij}-\bar{X}_{i\circ}-\bar{X}_{\circ j}+\bar{X})^2 , \quad \nu = (r-1)(s-1)$$

where

$$\bar{X}_{\circ j} = \frac{1}{r}\sum_i X_{ij} , \quad \bar{X}_{i\circ} = \frac{1}{s}\sum_j X_{ij} \quad \text{and} \quad \bar{X} = \frac{1}{rs}\sum_{i,j} X_{ij} .$$

Here we could be interested in either the row effects α_i, in which case $\theta_i = \mu+\alpha_i$, $\hat{\theta}_i = \bar{X}_{i\circ}$ and $b_i = s^{-1}$, or in the column effects β_j, in which case $\theta_j = \mu+\beta_j$, $\hat{\theta}_j = \bar{X}_{\circ j}$ and $b_i = r^{-1}$.

(b) With interaction.

Model: X_{ijg}, $i=1,\ldots,r$, $j=1,\ldots,s$, $g=1,\ldots,m$ are independent, normally distributed.

$$EX_{ijg} = \mu+\alpha_i+\beta_j+\gamma_{ij} \quad \text{where} \quad \sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0 .$$

$$\operatorname{var} X_{ijg} = \sigma^2$$

$$S^2 = \frac{1}{rs(m-1)} \sum_{i,j,g} (X_{ijg}-\bar{X}_{ij})^2 , \quad \nu = rs(m-1) .$$

where $\bar{X}_{ij} = \frac{1}{m} \sum_g X_{ijg}$.

Also in this case we would consider either row or column effects. For the row effects we would as in a) put $\theta_i = \mu+\alpha_i$ and $\hat{\theta}_i = \bar{X}_{i\circ}$, but $b_i = 1/sm$. For the column effects we have similarily $\theta_j = \mu+\beta_j$, $\hat{\theta}_j = \bar{X}_{\circ j}$ and $b = 1/rm$.

Here $\bar{X}_{i\circ} = \frac{1}{s\cdot m} \sum_{j,g} X_{ijg}$ and $\bar{X}_{\circ j} = \frac{1}{rm} \sum_{i,g} X_{ijg}$

We now return to the general model. It is well known that the S-method in this case is to state $\sum_i c_i\theta_i > 0$ with $\sum_i c_i = 0$ if

$$\sum_i c_i\hat{\theta}_i > \sqrt{(k-1)F}\, S \sqrt{\sum_i b_i c_i^2} \qquad (8)$$

where F is the upper ϵ-fractile in the F-distribution with $k-1$ and ν degrees of freedom. The probability of at least one false statement is then always less than or equal to ϵ. Now, the F-test for the homogenity hypothesis $H : \theta_1=\ldots=\theta_k$, is to reject H when

$$V = \frac{\sum b_i^{-1} (\hat{\theta}_i-\overset{*}{\theta})^2}{(k-1)S^2} > F \text{ , where } \overset{*}{\theta} = \frac{\sum b_i^{-1}\hat{\theta}_i}{\sum b_i^{-1}}$$

The test has level ϵ since V is F-distributed with $k-1$ and ν degrees of freedom when $\theta_1=\ldots=\theta_k$.

Now let $W_i = b_iS^2$, giving $\bar{\theta} = \frac{\sum W_i^{-1}\hat{\theta}_i}{\sum W_i^{-1}} = \frac{\sum b_i^{-1}\hat{\theta}_i}{\sum b_i^{-1}} = \theta^*$

and

$$\hat{\sigma}^2_{c^*\hat{\theta}} = S^2 \sum_{i=1}^{k} b_i c_i^2 \qquad (9)$$

from (7).

The equivalence of the F-test with the S-method follows now directly from theorem 1, which gives

$$\sum_{i=1}^{k} (\hat{\theta}_i-\hat{\theta})^2 b_i^{-1}S^{-2} \leq z \iff \sum_i c_i\hat{\theta}_i \leq \sqrt{z}\, S \sqrt{\sum_i b_i c_i^2} \quad \forall c \text{ , } \sum_i c_i = 0$$

With $z = (k-1)F$, we get

$$V \leq F \iff \sum_i c_i\hat{\theta}_i \leq \sqrt{(k-1)F}\, S \sqrt{\sum_i b_i c_i^2} \quad , \quad \forall c \text{ , } \sum_i c_i = 0 \qquad (10)$$

Now, the result (10) is well known. In the next section, however, we will consider the asymptotic situation. Here only the asymptotic distribution of the estimators is needed. A main theorem will be given.

Thereafter, in section 5 and 6, we apply the theorem to contingency tables and the non-parametric one-way layout with unequal, unknown group variances. In these situations we find interesting methods for linear contrasts which turn out to be natural modifications of the methods proposed for general linear functions by the author, [2], for contingency tables, and by Spjøtvoll, [5], for the one-way layout. At last we consider in section 7 the comparison of binomial probabilities.

4. The asymptotic case.

As in section 2 we let $\hat{\theta}_1,\ldots,\hat{\theta}_k$ be independent estimators of $\theta_1,\ldots,\theta_k$ respectively, $\hat{\theta}_i$ is based on n_i observations. Let $n = \sum_{i=1}^{k} n_i$. Throughout this section we assume that the following four assumptions hold.

A I. $n_i/n = \pi_i > 0$ is constant as n tends to infinity.

A II. The asymptotic distribution of $\sqrt{n_i}(\hat{\theta}_i-\theta_i)$ is $N(0,\sigma_i)$, denoted by $\sqrt{n_i}(\hat{\theta}_i-\theta_i) \xrightarrow{\mathcal{D}} N(0,\sigma_i)$.
$N(0,\sigma_i)$ denotes the normal distribution with mean zero and variance equal to σ_i^2.

A III. $W_1,\ldots,W_k$ are independent random variables satisfying $n_i W_i \xrightarrow{P} \sigma_i^2$, i.e. $n_i W_i$ is a consistent estimator of the asymptotic variance σ_i^2.

A IV. W_i and $\hat{\theta}_j$ are independent for $i \neq j$.

Then we have the following result.

LEMMA 2. Let $\bar{\theta}$ be defined by (4). Then the asymptotic distribution of $U = \sum_i W_i^{-1} (\hat{\theta}_i - \bar{\theta})^2$ is chi-square with $(k-1)$ degrees of freedom when $\theta_1 = \ldots = \theta_k$. We denote this by

$$U \overset{\mathcal{L}}{\sim} \chi^2(k-1) \text{ under the hypothesis (1).}$$

Proof. From Rao, [4], result 6a .2(v) the asymptotic distribution of U follows directly. The only difference here is that $n_i W_i$ is not a continuous function of $\hat{\theta}_i$. What we need, however, is that $n_i W_i \xrightarrow{P} \sigma_i^2$ which is true by A III, and hence the result follows by letting S_i^2 in Rao's result be equal to $n_i W_i$.

Q.E.D.

Let $z = z(k-1, \epsilon)$ be the upper ϵ-fractile in the chi-square distribution with $k-1$ degrees of freedom. Then a test for the hypothesis (1) is to reject H when $U > z$. We propose now the following method for linear contrasts.

$$\text{State } \sum_i c_i \theta_i > 0 \quad \text{if} \quad \sum_i c_i \hat{\theta}_i > \sqrt{z}\, \hat{\sigma}_{c,\hat{\theta}} \qquad (11)$$

where $\hat{\sigma}_{c,\hat{\theta}}$ is given by (7).

Then from theorem 1 we see that

U is not significant if and only if we can find no significant contrasts. Let $\epsilon(\theta_1, \ldots, \theta_k)$ be the probability of at least one false statement when $(\theta_1, \ldots, \theta_k)$ is the true values of the parameters. Then we have already shown, using lemma 2 and theorem 1, that $\lim_{n \to \infty} \epsilon(\theta, \ldots, \theta) = \epsilon$. The next result shows that ϵ is an upper bound for limsup $\epsilon(\theta_1, \ldots, \theta_k)$ for any set of values $\theta_1, \ldots, \theta_k$.

The next theorem also state simultaneous confidence intervals for all contrasts

$$\sum_i c_i\theta_i \ , \ \sum_i c_i = 0 \ .$$

THEOREM 2. Let (11) be the comparison procedure for linear contrasts. Then $\limsup_n \epsilon(\theta_1,\ldots,\theta_k) \le \epsilon$ for all $(\theta_1,\ldots,\theta_k)$. (12)
Simultaneous confidence intervals for $\sum_i c_i\theta_i$; $\sum_i c_i = 0$, are given by

$$\lim_{n\to\infty} P\Big(\sum_i c_i\hat{\theta}_i - \sqrt{z(k-1,\epsilon)}\,\hat{\sigma}_{c,\hat{\theta}} \le \sum_i c_i\theta_i \le \sum_i c_i\hat{\theta}_i + \sqrt{z(k-1,\epsilon)}\,\hat{\sigma}_{c,\hat{\theta}} \ \forall c, \sum_i c_i = 0\Big) = 1-\epsilon \ . \quad (13)$$

Proof.

$$\epsilon(\theta_1,\ldots,\theta_k) = P\Big(\bigcup_{\substack{\sum c_i\theta_i \le 0 \\ \sum c_i = 0}} \sum c_i\hat{\theta}_i > \sqrt{z_o}\sqrt{\sum W_i c_i^2}\Big) \ ,$$

where $z_o = z(k-1,\epsilon)$.

Let $Z_i = \dfrac{\hat{\theta}_i - \theta_i}{\sqrt{W_i}}$; $Z_i \xrightarrow{\mathcal{D}} X_i \sim N(0,1)$.

$Z_1,\ldots,Z_k$ are independent. Then

$\sum_i c_i\hat{\theta}_i = \sum_i c_i\sqrt{W_i}Z_i + \sum_i c_i\theta_i$. Let $h_i = c_i\sqrt{W_i}$. Then

$$\epsilon(\theta_1,\ldots,\theta_k) = P\Big(\bigcup_{\substack{\sum c_i\theta_i \le 0 \\ \sum c_i = 0}} \sum h_i Z_i + \sum c_i\theta_i > \sqrt{z_o}\hat{\sigma}_{c,\hat{\theta}}\Big)$$

$$= P\Big(\bigcup_{\substack{\sum c_i\theta_i \le 0 \\ \sum c_i = 0}} \sum_i (\sqrt{n}h_i)Z_i + \sqrt{n}\sum c_i\theta_i > \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}}\Big)$$

$$\le P\Big(\bigcup_{\substack{\sum c_i\theta_i \le 0, \\ \sum c_i = 0}} \sum_i (\sqrt{n}h_i)Z_i > \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}}\Big)$$

$$\le P\Big(\bigcup_{\sum c_i = 0} \sum_i (\sqrt{n}h_i)\,Z_i > \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}}\Big)$$

Now let $\hat{g}_i = \sqrt{n}\,h_i$, for $i = 1,\ldots,k$. $\hat{g}_i \xrightarrow{P} c_i\sigma_i/\sqrt{\pi_i}$.

Let $\hat{g} = (\hat{g}_1,\ldots,\hat{g}_k)'$ and $Z = (Z_1,\ldots,Z_k)'$. Then

$$\limsup_n\, \epsilon(\theta_1,\ldots,\theta_k) \le \limsup_n\, P\Big(\bigcup_{\sum c_i = 0} \hat{g}'Z > \sqrt{z_o}\,\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}}\Big)$$

Let $A = \bigcup_{\sum c_i = 0} (\hat{g}'Z > \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}})$. The complement of A, A^c, is then

$$A^c = \bigcap_{\sum c_i = 0} (\hat{g}'Z \le \sqrt{z_o}\,\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}})$$

Let $(z,w) = (z_1,\ldots,z_k,w_1,\ldots,w_k)$ be observed value of $Z_1,\ldots,Z_k$, $W_1,\ldots,W_k$. For observed (z,w) we let $g_i = \sqrt{n}\,c_i\sqrt{w_i}$ and $g = (g_1,\ldots,g_k)'$. I.e. g is the observed value of $\hat{g}$. Then

$$A^c = \{(z,w): g'z \le \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}} \quad \forall c, \textstyle\sum_i c_i = 0\}.$$

Consider for a given (z,w) the following subspace of $\mathbb{R}^k$:

$$V_{k-1}(w) = \{g:\ g_i = \sqrt{n}\,c_i\sqrt{w_i}\ ;\ \textstyle\sum_i c_i = 0\}.$$

As in the proof of theorem 1 we see that $d(V_{k-1}(w)) = k-1$. A^c can now be expressed as follows.

$$A^c = \{(z,w): g'z \leq \sqrt{z_o}\sqrt{g'g} \quad , \forall g \in V_{k-1}(w)\}.$$

Let $G_1,\ldots,G_{k-1}$ be an orthomormal basis for $V_{k-1}(w)$. i.e. $G_i'G_i = 1$ and $G_i'G_j = 0$ for $i \neq j$. Hence

$$V_{k-1}(w) = \{g \in \mathbb{R}^k : g = \sum_{i=1}^{k-1} \beta_i G_i \quad , \quad \forall \beta=(\beta_1,\ldots,\beta_{k-1})' \in \mathbb{R}^{k-1}\}.$$

This gives

$$(z,w) \in A^c \iff (\sum_{i=1}^{k-1} \beta_i G_i)'z \leq \sqrt{z_o}\sqrt{(\sum_{i=1}^{k-1} \beta_i G_i)'(\sum_{i=1}^{k-1} \beta_i G_i)} \quad \forall \beta \in \mathbb{R}^{k-1}$$

$$\iff \sum_{i=1}^{k-1} \beta_i (G_i'z) \leq \sqrt{z_o}\sqrt{\beta'\beta} \quad \forall \beta \in \mathbb{R}^{k-1} .$$

Now, let Y_i be the random variable given by the transformation

$y_i = G_i'z$ for $i = 1,\ldots,k-1$. Then

$$(z,w) \in A^c \iff \sum_{i=1}^{k-1} \beta_i y_i \leq \sqrt{z_o}\sqrt{\beta'\beta} \quad \forall \beta \in \mathbb{R}^{k-1} .$$

As in the proof of lemma 1 we see that this is equivalent with

$\sum_{i=1}^{k-1} y_i^2 \leq \sqrt{z_o}$. Hence

$A^c = \{Y'Y \leq z_o\}$, where

$Y' = (Y_1,\ldots,Y_{k-1})$.

Now we need a general result from asymptotic distribution theory. It states that if $X_n = (X_{1n},\ldots X_{kn})' \xrightarrow{\mathcal{D}} X = (X_1,\ldots X_k)'$ and f is a continuous function from $\mathbb{R}^k$ to $\mathbb{R}^m$ then $f(X_n) \xrightarrow{\mathcal{D}} f(X)$.

Let us first consider the asymptotic distribution of Y given $W = w$. Then $Y = G'Z$ is a continuous function of Z , where $G = (G_1,\ldots,G_{k-1})$ and $G_i = G_i(w)$, $i=1,\ldots k-1$ is an orthonormal basis of $V_{k-1}(w)$. Now $Z \xrightarrow{\mathcal{D}} X =(X_1,\ldots,X_k)' \sim N_k(0,I)$. Hence the conditional distribution of Y given $W=w$ converges to the distribution of $V = G'X$.
Let $V = (V_1,\ldots,V_{k-1})$. Then $V_i = G_i'X \sim N(0,1)$ for $i=1,\ldots,k-1$ and $\operatorname{cov}(V_i,V_j) = 0$, i.e. $V_1,\ldots,V_{k-1}$ are independent and $N(0,1)$. This imply that the conditional distribution of Y given $W=w$ converges to $N_{k-1}(0,I)$ which is independent of w . This means that

$$Y \xrightarrow{\mathcal{D}} V \sim N_{k-1}(0,I) .$$

Now $Y'Y = \sum_{i=1}^{k-1} Y_i^2$ is a continuous function of Y implying that

$$Y'Y \xrightarrow{\mathcal{D}} V'V \sim \chi^2(k-1) . \qquad (14)$$

From (14) we get

$$\limsup_n \varepsilon(\theta) \le \limsup_n (A) = 1 - \lim_n P(A^c) = 1 - \lim_n P(Y'Y \le z) = \varepsilon ,$$

and (12) is shown.

To construct simultaneous confidence intervals for $\sum_i c_i\theta_i$ with $\sum_i c_i = 0$ we make use of the fact that

$$A^c = \bigcap_{\sum_i c_i = 0} (g'Z \le \sqrt{z_o}\sqrt{n}\,\hat{\sigma}_{c,\hat{\theta}}) \quad \text{and} \quad \lim_n P(A^c) = 1-\varepsilon .$$

$$g'Z = \sum_{i=1}^{k} g_i Z_i = \sum_i \sqrt{n}\, c_i \sqrt{W_i}\, \frac{(\hat{\theta}_i - \theta_i)}{\sqrt{W_i}} = \sqrt{n} \sum_i c_i (\hat{\theta}_i - \theta_i) \ .$$

This gives $$A^c = \bigcap_{\sum c_i = 0} \left(\sum_i c_i (\hat{\theta}_i - \theta_i) \leq \sqrt{z_o}\ \hat{\sigma}_{c,\hat{\theta}} \right) .$$

Now it is obvious that

$$\sum_i c_i (\hat{\theta}_i - \theta_i) \leq \sqrt{z_o}\ \hat{\sigma}_{c,\hat{\theta}} \quad \text{for all} \quad c \quad \text{with} \quad \sum_i c_i = 0$$

$$\Updownarrow$$

$$\left| \sum_i c_i (\hat{\theta}_i - \theta_i) \right| \leq \sqrt{z_o}\ \hat{\sigma}_{c,\hat{\theta}} \quad \text{for all} \quad c \quad \text{with} \quad \sum_i c_i = 0 \ .$$

Hence

$$\lim_{n \to \infty} P \left(\bigcap_{\sum c_i = 0} \left| \sum_i c_i (\hat{\theta}_i - \theta_i) \right| \leq \sqrt{z_o}\ \hat{\sigma}_{c,\hat{\theta}} \right) = 1 - \epsilon \ , \quad \text{and} \quad (13) \quad \text{is}$$

proved. Q.E.D.

The equivalence between the chi-square test of homogenity and the multiple comparison procedure (11) provides us with a useful tool for judging contrasts. Accordingly we should always first find out if U is significant. If we don't reject the homogenity hypothesis we would not have to bother about looking for significant contrasts, because we know we will not find any.
On the other hand, if U is significant we know there are (14)
significant contrasts too. Notice that if we want to test a specific contrast, i.e. a specific hypothesis

$$H_c : \sum_{i=1}^{k} c_i \theta_i = 0 \tag{15}$$

the test will be to state $\sum_{i=1}^{k} c_i \theta_i > 0$ if

$$\sum_i c_i \hat{\theta}_i > x(\epsilon)\ \hat{\sigma}_{c,\hat{\theta}} \tag{16}$$

where $x(\epsilon)$ is the upper ϵ-fractile in the $N(0,1)$.

The test has asymptotic level ϵ, since

$$\frac{\sum c_i(\hat{\theta}_i-\theta_i)}{\hat{\sigma}_{c,\hat{\theta}}} = \frac{\sqrt{n}\sum c_i(\hat{\theta}_i-\theta_i)}{\sqrt{n}\hat{\sigma}_{c,\hat{\theta}}} \xrightarrow{\mathcal{D}} N(0,1) . \qquad (17)$$

So the method (11) for contrasts is essentially the same as (16). The only thing we have to do is to adjust the critical level from $x(\epsilon)$ to $\sqrt{z(k-1,\epsilon)}$. As mentioned above the rule for judging contrasts should be used as described below.

Multiple test for contrasts.

Step 1. Let $H : \theta_1=\ldots=\theta_k$.

Test for H : Reject if $U = \sum_i W_i^{-1}(\hat{\theta}_i-\bar{\theta})^2 > z(k-1,\epsilon)$

If we do reject we continue to step 2.

If $U \leq z(k-1,\epsilon)$ we stop.

Step 2.

State $\sum_i c_i\theta_i > 0$, $\sum_i c_i = 0$ if

$\sum_i c_i\hat{\theta}_i > \sqrt{z(k-1,\epsilon)}\,\hat{\sigma}_{c,\hat{\theta}}$.

We now go on to discuss some applications of this method.

5. Comparison of independent contingency tables.

The situation is described in [2]. We give a brief abstract of it. We want to compare k independent two-way contingency tables by using a measure of association. For a presentation of measures of association we refer to [1], part 1 or [3].

Let p_{ijr} , $i=1,\ldots,v_r$ and $j=1,\ldots,w_r$ be the cell-probabilities in table r , for $r=1,\ldots,k$. I.e. we have $\nu_r=v_r \cdot w_r$ cells in table r . X_{ijr} is the number of observations in cell (i,j) and n_r is the total number of observations in table r . The relative frequencies are denoted by $q_{ijr} = X_{ijr}/n_r$. Let $n=\sum_{r=1}^{k} n_r$ and $\pi_r = n_r/n$. The π_r's are considered as constants as n tends to infinity.

We now let

$$p_r = (p_{11r},\ldots,p_{v_r 1,r},\ldots,p_{v_r,w_r,r})$$

$$q_r = (q_{11r},\ldots,q_{v_r 1,r},\ldots,q_{v_r,w_r,r})$$

for $r=1,\ldots,k$. We assume that the tables are independent, i.e. $q_1,\ldots,q_k$ are independent random variables.

Let $p=(p_1,\ldots,p_k)$ and $q=(q_1,\ldots,q_k)$.

All $\{p_{ijr}\}$ are assumed to be positive.

Assume we have chosen d as a suitable measure of association. d is assumed to have continuous partial derivatives as a function of the cell-probabilities. We let d_i be the measure d in table i , for $i=1,\ldots,k$. Then $d_i=d_i(p_i)$ is a function of ν_i variables with continuous partial derivatives. We want to consider linear contrasts in the d_i's , $\sum_{i=1}^{k} c_i d_i$ with $\sum_i c_i = 0$.

Consistent estimators $\hat{d}_i$ are obtained by letting

$$\hat{d}_i = d_i(q_i) \quad , \quad \text{for } i=1,\ldots,k$$

Then $\hat{d}_1,\ldots,\hat{d}_k$ are independent by assumption.

From the main theorem in [1], part 1 we have that

$$\sqrt{n_i}(\hat{d}_i - d_i) \xrightarrow{\mathcal{D}} N(0, \sigma_{d,i}) \tag{18}$$

where

$$\sigma^2_{d,r} = \sum_{i=1}^{v_r} \sum_{j=1}^{w_r} p_{ijr} \, (d_{ij,r} - d^*_r)^2 \text{ , provided } \sigma^2_{d,r} > 0 \text{ .} \tag{19}$$

Here

$$d_{ij,r} = \partial d_r / \partial p_{ijr}$$

and

$$d_r^* = \sum_{i,j} d_{ij,r} \cdot p_{ijr} \text{ .}$$

Consistent estimators of the asymtotic variances $\sigma^2_{d,r}$ of $\sqrt{n_r}\hat{d}_r$ are given by

$$S^2_{d,r} = \sum_i \sum_j q_{ijr} (\hat{d}_{ij,r} - \hat{d}^*_r)^2 \tag{20}$$

where

$$\hat{d}_{ij,r} = d_{ij,r}(q_r)$$

and

$$\hat{d}^*_r = \sum_i \sum_j q_{ijr} \, \hat{d}_{ijr}$$

I.e. $S^2_{d,r} = \sigma^2_{d,r}(q_r)$.

An estimator of the variance of $c'\hat{d} = \sum_i c_i \hat{d}_i$ is given by

$$\hat{\sigma}^2_{c'\hat{d}} = \sum_{i=1}^{k} \frac{c_i^2 \, S^2_{d,i}}{n_i} \tag{21}$$

In [2] the author gave a method for general linear functions as follows. State $\sum_i c_i d_i > 0$ if

$$\sum_i c_i \hat{d}_i > \sqrt{z(k,\varepsilon)} \, \hat{\sigma}_{c'\hat{d}} \tag{22}$$

It is readily seen from theorem 3 in [2] that

$$\limsup_{n\to\infty} P\ (\text{at least one false statement} : \sum_i c_i d_i > 0) \leq \epsilon$$

for all values of $d_1,\ldots,d_k$.

By letting $W_i = S^2_{d,i}/n_i$ we see that the assumptions AI - AIV in section 4 are fulfilled. We are now interested in modifying the rule (22) to linear contrasts. From (11), the only thing we have to do then is to substitute $z(k,\epsilon)$ with $z(k-1,\epsilon)$ which is the natural modification we would expect. Hence for all contrasts we state $\sum_i c_i d_i > 0$ if

$$\sum_i c_i \hat{d}_i > \sqrt{z(k-1,\epsilon)}\ \hat{\sigma}_{c,\hat{d}} \tag{23}$$

To be exact the method consists of two steps:

Step 1. Test the hypothesis

$$H : d_1 = \ldots = d_k$$

$$\text{Reject if } U = \sum_i \frac{n_i(\hat{d}_i - \bar{d})^2}{S^2_{d,i}} > z(k-1,\epsilon)$$

$$\text{where } \bar{d} = \Big(\sum_i \frac{n_i}{S^2_{d,i}}\Big)^{-1} \sum_i \frac{n_i \hat{d}_i}{S^2_{d,i}} \quad .$$

If we reject we go to step 2, if not the procedure stops.

Step 2.

State $\sum_i c_i d_i > 0$ if

$$\sum_i c_i \hat{d}_i > \sqrt{z(k-1,\epsilon)}\ \hat{\sigma}_{c,\hat{d}} \quad .$$

From theorem 2 : $\limsup_{n\to\infty} P\ (\text{at least one false statement}) \leq \epsilon$.

Simultaneous confidence intervals for all linear contrasts in the measures $d_1,\ldots,d_k$ are given by:

$$\sum_i c_i d_i \in \left[\sum_i c_i \hat{d}_i \pm \sqrt{z(k-1,\epsilon)}\,\hat{\sigma}_c, \hat{d}\right] \qquad (24)$$

6. One-way layout with unequal group variances. Non-parametric case.

Spjøtvoll proposed in [5] a method for all linear functions of the means in the one-way layout with unequal group variances, for the normal model. We consider now a more general situation, not assuming normality, and derive from theorem 2 a method for testing all linear contrasts of the means. This method will asymptotically be the natural modification of Spjøtvoll's procedure for all linear functions. Our model is as follows.

Let the random variables $X_{ij}, i = 1,\ldots,k$, $j=1,\ldots,n_i$ be independent with

$$EX_{ij} = \mu_i \text{ and } \operatorname{var} X_{ij} = \sigma_i^2 .$$

For each i, $X_{i1},\ldots,X_{i,n_i}$ are identically distributed with finite third and fourth order moments, i.e. EX_{ij}^3 and EX_{ij}^4 are finite .

Let $n = \sum_{i=1}^{k} n_i$. We assume $\pi_i = n_i/n > 0$ and constant as $n \to \infty$.

Let

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \quad , \quad \text{for } i=1,\ldots,k .$$

$\hat{\mu}_1,\ldots,\hat{\mu}_k$ are independent and

$$\sqrt{n_i}(\hat{\mu}_i - \mu_i) \xrightarrow{\mathcal{D}} N(0,\sigma_i) ,$$

from the central limit theorem for independent, identically distributed random variables.

Hence the assumptions A I and A II in section 4 are satisfied. We have left to find a consistent estimator for σ_i^2 , $i=1,\ldots,k$. Let now

$$S_i^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (X_{ij}-\bar{X}_i)^2 . \tag{25}$$

S_i^2 and $\hat{\mu}_j$ are of course independent for $i \neq j$.

$$ES_i^2 = \sigma_i^2 \quad , \quad \text{for} \quad i=1,\ldots,k .$$

We shall show that S_i^2 is a consistent estimator of σ_i^2 , i.e. that var $S_i^2 \to 0$ as $n \to \infty$.

$$\text{var}\,(S_i^2) = ES_i^4 - \sigma_i^4 \quad , \text{ and}$$

$$ES_i^4 = (n_i-1)^{-2} \sum_{j \neq l} E(X_{ij}-\bar{X}_i)^2(X_{il}-\bar{X}_i)^2 + (n_i-1)^{-2} \sum_{j} E(X_{ij}-\bar{X}_i)^4$$

$E(X_{ij}-\bar{X}_i)^2(X_{il}-\bar{X}_i)^2 = \rho_i$, independent of $j \neq l$.

$E(X_{ij}-\bar{X}_i)^4 = \xi_i$, independent of j . Hence

$$ES_i^4 = \frac{n_i}{n_i-1} \rho_i + \frac{n_i}{(n_i-1)^2} \xi_i .$$

After some calculations we find that

$$\rho_i \underset{n \to \infty}{\to} (\sigma_i+\mu_i^2)^2 - 2\mu_i^2 (\sigma_i^2+\mu_i^2) + \mu_i^4 = \sigma_i^4$$

It is readily seen that $\frac{n_i}{(n_i-1)^2} \xi_i \to 0$ as $n \to \infty$.

This gives that

$$ES_i^4 \to \sigma_i^4 \quad \text{and} \quad \text{var}(S_i^2) \to 0 \quad \text{as} \quad n \to \infty .$$

I.e. S_i^2 is unbiased and consistent estimator of σ_i^2 .

Let $W_i = S_i^2/n_i$, and we see that also conditions A III and A IV are fulfilled. We shall now consider linear contrasts,

$\sum_i c_i\mu_i$ with $\sum_i c_i = 0$.

An unbiased estimator of $\operatorname{var}(\sum_i c_i\hat{\mu}_i)$ is given by

$$\hat{\sigma}^2_{c'\hat{\mu}} = \sum_{i=1}^{k} \frac{c_i^2 S_i^2}{n_i} \tag{26}$$

From (11) we have the following test for linear contrasts:

State $\sum_i c_i\mu_i > 0$ if

$$\sum_i c_i\hat{\mu}_i > \sqrt{z(k-1,\epsilon)}\ \hat{\sigma}_{c'\hat{\mu}} \tag{27}$$

Let $$U = \sum_i \frac{n_i(\hat{\mu}_i-\bar{\mu})^2}{S_i^2}$$ where

$$\bar{\mu} = (\sum_i \frac{n_i}{S_i^2})^{-1} \sum_i \frac{n_i\hat{\mu}_i}{S_i^2} \quad .$$

From lemma 2:

$$U \xrightarrow{\mathcal{D}} \chi^2(k-1) \quad \text{when} \quad \mu_1=\ldots=\mu_k$$

A test for the hypothesis $H : \mu_1=\ldots=\mu_k$ is then to reject if

$$U > z(k-1,\epsilon) \ . \tag{28}$$

From theorem 1, the test-criterion (28) is equivalent with stating at least one contrast $\sum_i c_i\mu_i > 0$ with the procedure (27) . From theorem 2 we have

$$\limsup_n P \text{ (at least one false statement)} \leq \epsilon \ ,$$

for all values of $\mu_1,\ldots,\mu_k$.

The asymptotic distribution of $\sum_i n_i S_i^{-2} (\hat{\theta}_i - \theta_i)^2$ is chi-square with k degrees of freedom. It follows that Spjøtvoll's method for general linear functions in this asymptotic case would be to state $\sum_i c_i \theta_i > 0$ if

$$\sum_i c_i \hat{\theta}_i > \sqrt{z(k,\epsilon)}\, \hat{\sigma}_{c,\hat{\mu}} \tag{29}$$

I.e. $z(k,\epsilon)$ replaces A^2 given in [5]. We see that our method (27) for contrasts is the natural modification of (29).

Simultaneous confidence intervals for all linear contrasts are given by

$$\sum_i c_i \theta_i \in \left[\sum_i c_i \hat{\theta}_i \pm \sqrt{z(k-1,\epsilon)}\, \hat{\sigma}_{c,\hat{\mu}}\right] \tag{30}$$

It should be noticed that the non-parametric model in this section covers the usual normal-model for unequal group variances, so that the method for contrasts proposed in (27) are asymptotically correct also in the normal model.

As a last application we consider comparison of k independent binomial sequences.

7. Comparison of binomial distributions.

The model is as follows. $X_1,\ldots,X_k$ are independent. X_i is binomially distributed with parameters (n_i,p_i), i.e.

$$P(X_i=x) = \binom{n_i}{x} p_i^x (1-p_i)^{n_i-x} \quad \text{for} \quad x=0,\ldots,n_i .$$

p_i is assumed to be positive for all i. We are interested in contrasts $\sum_{i=1}^{k} c_i p_i$, $\sum_i c_i = 0$, and also in testing the hypothesis

$$H : p_1 = \ldots = p_k \tag{31}$$

Let $\pi_i = n_i/n$, where $n = \sum_i n_i$. The π_i's are assumed to be positive constants as n tends to infinites . Let $\hat{p}_i = X_i/n_i$. Then

$$\sqrt{n_i}(\hat{p}_i - p_i) \xrightarrow{\mathcal{D}} N(0,\sigma_i) \quad , \quad \text{where} \quad \sigma_i^2 = p_i(1-p_i) \ .$$

A consistent estimator of σ_i^2 is

$$S_i^2 = \hat{p}_i(1-\hat{p}_i) \ . \quad \text{Let then} \quad W_i = S_i^2/n_i \ .$$

Then all the assumptions A I - A IV are satisfied. From lemma 2 we find that

$$Z = \sum_i \frac{n_i(\hat{p}_i - \bar{p})^2}{\hat{p}_i(1-\hat{p}_i)} \tag{32}$$

is asymptotically $\chi^2(k-1)$ when $p_1 = \ldots = p_k$. Here

$$\bar{p} = \sum_i \frac{\hat{p}_i n_i}{\hat{p}(1-\hat{p}_i)} \Big/ \sum_i \frac{n_i}{\hat{p}_i(1-\hat{p}_i)} \tag{33}$$

is the modified chi-square minimum estimator, since $\bar{p}$ minimizes

$$Q = \sum_i \frac{n_i(\hat{p}_i - p)^2}{\hat{p}_i(1-\hat{p}_i)} = \sum_i \left[\frac{(X_i - n_i p)^2}{X_i} + \frac{(n_i - X_i - n_i(1-p))^2}{n_i - X_i} \right] .$$

The test for the hypothesis (31) is then to reject H if

$$Z > z(k-1,\epsilon) \tag{34}$$

This test is well known as one of the usual chi-square tests for (31) .

An estimator of the variance of $c'\hat{p} = \sum_i c_i \hat{p}_i$ is given by

$$\hat{\sigma}^2_{c'\hat{p}} = \sum_i c_i^2 \frac{\hat{p}_i(1-\hat{p}_i)}{n_i} \tag{35}$$

From (11) we have now the following test for linear contrasts.

$$\text{State } \sum_i c_i p_i > 0 \quad \text{if} \quad \sum_i c_i \hat{p}_i > \sqrt{z(k-1,\epsilon)}\, \hat{\sigma}_{c'\hat{p}} \tag{36}$$

Now, from theorem 1 the chi-square test-criterion (34) is equivalent with stating at least one false statement with test (36) . Simultaneous confidence intervals for linear contrasts are given by

$$\sum_i c_i p_i \in \left[\sum_i c_i \hat{p}_i \pm \sqrt{z(k-1,\epsilon)}\, \hat{\sigma}_{c'\hat{p}}\right] \tag{37}$$

References.

[1] Bjørnstad, J.F. (1975): "Inference theory in contingency tables", Statistical Research Report, Univ. of Oslo.

[2] Bjørnstad, J.F. (1975): "Comparison of contingency tables", Statistical Research report, Univ. of Oslo.

[3] Goodman, L.A. & Kruskal, W.H. (1954): "Measures of association for cross classifications", J.Am.Statist.Ass. Vol.4, 732-764.

[4] Rao, C.R. (1965): "Linear Statistical Inference and Its Applications", John Wiley & Sons,Inc.

[5] Spjøtvoll, E. (1972): "Joint confidence intervals for all linear functions of means in the one-way layout with unknown group variances". Biometrika, Vol. 59, 683-685.